RECEIVED

FEB 24 2006

Technology Center 2600

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Appl. No.	: 09/759,486
Appellant(s)	: PELLETIER, Daniel
Filed	: 12 January 2001
Title	: METHOD AND APPARATUS FOR DETERMINING CAMERA MOVEMENT CONTROL CRITERIA
TC/A.U.	: 2615
Examiner	: JONES, Heather R.
Atty. Docket	: US 010002

CERTIFICATE OF MAILING OR TRANSMISSION
I certify that this correspondence is being:

[✓] deposited with the U.S. Postal Service with sufficient postage as first-class mail in an envelope addressed to:

Board of Patent Appeals & Interferences
United States Patent & Trademark Off.
P.O. Box 1450
Alexandria, VA 22313-1450

[ ] transmitted by facsimile to the U.S. Patent and Trademark Office at 703-872-9318.

On: Feb. 15, 2006

By: John C Fox

APPELLANT'S APPEAL BRIEF
CORRECTED

RECEIVED

FEB 17 2006

BOARD OF PATENT APPEALS AND INTERFERENCES

Board of Patent Appeals and Interferences
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

This corrected Brief of Appellant is in response to a Notification of Non-Compliant Appeal Brief dated 20 January 2006. The non-compliant Brief followed a Notice of Appeal, dated 12 September 2005, appealing the decision dated 24 June 2005, of the Examiner finally rejecting claims 1, 3-7 and 9-19 of the application. All requisite fees set forth in 37 CFR 1.17(c) for this Brief are hereby authorized to be charged to Deposit Account No. 501,850.

REAL PARTY IN INTEREST

The real party in interest in this appeal is the assignee of all rights in and to the subject application, Koninklijke Philips Electronics, N.V. of The Netherlands.

RELATED APPEALS AND INTERFERENCES

To the best of the knowledge of the undersigned, no other appeals or interferences are known to Appellants, Appellants' legal representatives, or assignee which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

STATUS OF CLAIMS

Of the original claims 1-17, claims 1-15 were amended and claims 18 and 19 were added by amendment dated 6 July 2004, claims 1, 7, 18 and 19 were amended and claims 2 and 8 were cancelled by amendment dated 3 March 2005.

Claims 1, 3-7 and 9-19 now stand finally rejected as set forth in the final Office Action dated 24 June 2005, and are the subject of this appeal.

STATUS OF AMENDMENTS

No amendments were offered subsequent to the final Office action. All amendments have been entered.

SUMMARY OF THE CLAIMED SUBJECT MATTER

(Since the Specification does not contain line numbers, the following references to line numbers were determined by counting lines of text including section headings; blank lines were not counted, nor were tables or table labels.)

This invention relates to camera control, and to dynamically determining criteria to be used to control camera movement sequences based on the content of the scene being viewed. (Specification, page 1, lines 4-6)

The invention includes a method for automatically controlling the movements of at least one camera or camera lens to change the prospective of a scene viewed by said at least one camera or camera lens, the method comprising the steps of:

selecting at least one sequence of camera parametrics from a plurality of sequences of camera parametrics (150, 500), wherein said at least one sequence of camera parametrics (150, 500) is selected from the group of camera movements including scanning, zooming, tilting, orientating, panning, fading, zoom-and-pull-back, fade-in, fade-out, and wherein said parametrics provide instruction (170, 560) to control movement of said at least one camera or camera lens;

determining criteria for executing said selected sequence of camera parametrics (550), wherein said criteria are responsive to at least one high level parameter (140) of at least one object contained in said scene (510, 520); and

adjusting movement of said at least one camera or camera lens in response to said determined criteria (560). (Claim 1; Specification, page 5, lines 1-15; page 8, lines 12-14; page 9, lines 2-4; Figs. 1 and 3A)

The invention also includes an apparatus (200) for automatically controlling the movements of at least one camera or camera lens to change the prospective of a scene viewed by said at least one camera or camera lens, said apparatus comprising:

a processor (210) operative to:

receive a first input for selecting at least one sequence of camera parametrics from a plurality of sequences of camera parametrics, wherein said at least one sequence of camera parametrics is selected from the group of camera movements including scanning, zooming, tilting, orientating, panning, fading, zoom-and-pull-back, fade-in, fade-out, and wherein said parametrics provide instruction to control movement of said at least one camera or camera lens (220; 230; 240; 270);

receive a second input comprising at least one high level parameter of at least one object contained in said scene;

determine criteria for executing said selected sequence of camera parametrics, wherein said criteria are responsive to said at least one high level parameter; and

means for adjusting movement of said at least one camera or camera lens in response to said determined criteria. (Claim 7; Specification, page 3, lines 15, 16; page 7, lines 7 and 8; page 8, lines; page 9, line 19-21; page 10, lines 1-22; page 11, lines 2-19; Figs. 4A and 4B)

According to one embodiment of the method and apparatus of the invention, the at least one high level parameter (140) includes the number of objects within said scene (510). (Claims 3 and 9; Specification, page 5, lines 1-14; page 8, line 14; Figs. 1 and 3A)

According to another embodiment of the method and apparatus of the invention, the at least one high level parameter (140) includes the position of at least one object within said scene (520). (Claims 4 and 10; Specification, page 5, lines 1-14; page 8, line 14; Figs. 1 and 3A)

According to another embodiment of the method and apparatus of the invention, the at least one high level parameter (140) includes speech recognition of at least one object within said scene (120). (Claims 5 and 11; Specification, page 5, lines 1-14; Fig. 1)

According to another embodiment of the method and apparatus of the invention, the at least one high level parameter (140) includes an audio input of at least one object within said scene (130). (Claims 6 and 12; Specification, page 5, lines 1-14; Fig. 1)

According to another embodiment of the apparatus of the invention, the means for adjusting camera movement effects outputting of said criteria over a serial connection. (Claim 13)

According to another embodiment of the apparatus of the invention, the means for adjusting said camera movement effects outputting of said criteria over a parallel connection. (Claim 14)

According to another embodiment of the apparatus of the invention, the means for adjusting said camera movement effects outputting of said criteria over a network. (Claim 15)

According to another embodiment of the apparatus of the invention, the camera movement is accomplished electronically. (Claim 16)

According to another embodiment of the apparatus of the invention, the camera movement is accomplished mechanically. (Claim 17)

According to another embodiment of the method of the invention, the method includes:

- locating the at least one object in an image of the scene (520, 552);
- determining the object closest to a predetermined location in the image; and
- adjusting the movement of the at least one camera or camera lens in response to said determination. (Claim 18; Specification; page 8, line 14; page 9, line 10; Figs. 3A and 3B)

According to another embodiment of the method of the invention, the method includes:

- locating the at least one object in an image of the scene (520, 552);
- determining the object closest to the center of the image;
- determining the percentage of the scene around said closest object (554); and
- adjusting the zoom level of the at least one camera or camera lens in response to said percentage determination (560). (Claim 19; Specification; page 8, line 14; page 9, lines 4-11; Figs. 3A and 3B)

GROUND(S) OF REJECTION TO BE REVIEWED ON APPEAL

The grounds of rejection to be reviewed on appeal are:

1. Are claims 1, 3-7, 9-12 and 16-19 anticipated under 35 USC 102(e) by Chim (U.S. patent 6,275,258)?

2. Are claims 13-15 unpatentable under 35 USC 103(a) over Chim, as applied to claim 7 above, and further in view of Steinberg et al. (U.S. patent 6,750,902)(herein 'Steinberg')?

ARGUMENT

1. Are claims 1, 3-7, 9-12 and 16-19 anticipated under 35 USC 102(e) by Chim?

Claims 1, 3-7, 9-12 and 16-19 are rejected under 35 USC 102(e) as being anticipated by Chim.

Chim discloses a voice responsive image tracking system, which continuously tracks sound emitting objects by providing sound sensing means and a processor for directing a camera toward the sound source. See col. 3, line 36 through col. 4, line 3.

The relative signal levels of the sound sensing means, e.g., microphones, are continuously monitored for movement of the speaker for panning or zooming the camera, or both. See col. 4, lines 40-42.

Characteristics of audio signals are processed by an interface for determining movement of the speaker for directing the camera. As the characteristics sensed by the microphones change, the interface directs the camera toward the speaker. The interface continuously directs the camera, until the change

in the characteristics stabilizes, thus precisely directing the camera toward the speaker. See Abstract; col. 4, lines 43-58.

Thus, Chim's pan and zoom operations are governed by a single instruction, i.e., to find a speaker by panning and zooming the camera until the relative strengths of audio signals from a set of microphones are stabilized.

In contrast to the teachings of Chim, Appellant's claims 1 and 7 call for selecting at least one sequence of camera parametrics from a plurality of sequences of camera parametrics, including scanning, zooming, tilting, orientating, panning, fading, zoom-and-pull-back, fade-in and fade-out.

Since Chim does not disclose selecting at least one sequence of camera parametrics from a plurality of sequences of camera parametrics, Chim fails to anticipate the rejected claims, and it is urged that the rejection be reversed.

In response to Appellant's argument, the Examiner has responded that Chim discloses selecting at least two sequences of camera parametrics, panning and zooming, citing col. 4, lines 51-54 of the reference.

The cited passage states: 'The computer pans or zooms the camera toward the microphone transmitting the increasing signal level until the change in relative signal levels transmitted from the microphones stabilizes.'

The passage does not state or imply that panning and zooming are selected from a plurality of sequences of camera parametrics. In contrast, as already pointed out, panning and zooming are predetermined as the only types of camera movement to be employed, and the instructions for panning and zooming are also predetermined, i.e., to continue panning and zooming until the change in relative signal levels transmitted from the microphones stabilizes.

Thus, these instructions are not camera parametrics, i.e., these instructions are not generalized instructions for performing known camera movements. Moreover, these instructions are not selected from a plurality of sequences of camera parametrics.

The Examiner has argued that Appellant's claims do not call for each sequence to be a set of rules for determining camera movements. However, Appellant need not include the definition of a term in a claim, where the specification clearly sets forth that definition. The term 'sequence of camera parametrics' is defined in the specification as generalized instructions for performing known camera movements, at page 3, lines 10-13.

Examples of these sequences for zooming and panning are shown in Tables I and II, respectively, of the specification. Each sequence is more than just zooming or panning. Each sequence is a set of rules for determining the manner of execution of the zoom or pan operation.

Claims 1 and 7 require the selection of one or more sequences from a plurality of sequences.

In contrast, Chim does not teach or suggest selecting a sequence of camera parametrics from a plurality of such sequences. Chim merely teaches interface means for controlling camera movement (zooming or panning) in response to changes in the relative strength of audio signals from a set of microphones, until the changes in the audio signals are stabilized. See, e.g., col. 4, lines 51-54 of Chim.

Moreover, Appellant's claims require 'determining criteria for executing said selected sequence of camera parametrics', whereas Chim's criteria for camera movement is not determined, but rather has been predetermined, and is always the same,

i.e., the stabilization of the relative strength of audio signals from a set of microphones.

Thus, Chim does not teach or suggest 'determining criteria for executing said selected sequence of camera parametrics', as called for by Appellant's claims.

Regarding claims 3 and 9, Chim is not able to determine the number of objects in a scene. Chim only provides for determining the location of an object based on sounds detected from that object.

Thus, Chim states at col. 4, lines 63-67, that 'Using triangulation techniques and stereophonic microphones, the present invention provides a natural transition when tracking different speakers and is able to precisely determine **the position of each speaker when they are talking.**' (emphasis added).

In response to Appellant's argument that Chim is not able to determine the number of objects in a scene, the Examiner has stated that Chim discloses that his system can determine the current speaker from several different speakers, citing col. 4, lines 63-67. Thus, it is argued, this determination inherently includes the ability to determine the number of objects in a scene.

However, scenes include objects other than speakers, such as people who never speak and inanimate objects. Chim would not be able to locate these at all, since his system relies strictly on audio signals from speakers. Moreover, Chim doesn't even provide means for keeping track of the number of speakers. Speakers could come and go from the scene without Chim's system being aware, since the speakers are not uniquely identified, but merely tracked based on audio signal levels.

The Examiner has further stated that determining the positions for objects in a room go hand-in-hand with determining how many objects are in a room.

However, Chim does not keep track of how many speakers there are in a room. Chim rather continuously tracks the changing levels of audio signals in order to find the current speaker. Chim continuously moves from one speaker to the next, without any attempt to keep track of the number or location or identity of the speakers.

Regarding claims 5 and 11, Chim does not disclose speech recognition, but only audio detection via one or more microphones. Speech recognition is commonly understood to mean conversion of speech to digital signals, not audio signals.

In response to Appellant's argument that Chim does not disclose speech recognition, the Examiner has stated that audio detection of speech is the same as speech recognition.

However, audio detection is not the same as speech recognition. Chim only monitors relative signal levels. There is no teaching or suggestion of any effort to distinguish speech from any other sound. Moreover, there would be no need to do so. Consider the case of a speakerphone which is switched between transmit and receive states by so-called 'voice activation'. Such a system is activated by sound of any kind, not strictly by voice. Thus, a kick of the table or a rustling of papers can inadvertently switch the device. To provide actual voice recognition would involve a needless level of sophistication and expense.

Regarding claim 18, Chim does not disclose, literally or inherently, determining the object closest to a predetermined location in the image. Chim merely detects the position of an object based on calculating the position (e.g., by

triangulation) of an object based on the sound issuing from that object. Thus, the camera is instructed to pan to that location. There is no need, and indeed, Chim does not teach, to determine the distance of one object from another.

In response to Appellant's argument that Chim does not disclose determining the object closest to a predetermined location in the image, the Examiner has responded that in order to have the speakers captured in the center of the image, Chim would have to determine the object closest to a predetermined location or the object closest to the center of the image.

However, Chim controls camera movement in order to stabilize the changing audio levels from the microphones. This control is not the same as determining the object closest to a predetermined location or the object closest to the center of the image. Rather, this control finds the object which is emitting sound by triangulation of the audio signals from multiple strategically placed microphones. The sound-emitting object, i.e., speaker, need not be in a fixed location, but in fact may be moving about the room. See, e.g., col. 3, line 50.

Regarding claim 19, Chim does not disclose, literally or inherently, determining the object closest to the center of the image. Chim merely detects the position of an object based on calculating the position (e.g., by triangulation) of an object based on the sound issuing from that object. Thus, the camera is instructed to pan to that location. There is no need, and indeed, Chim does not teach, to determine the distance of one object from another.

With respect to claims 4, 6, 10, 12, 16 and 17, these claims are patentable, *inter alia*, by virtue of their dependency on claims 1 and 7.

For all of the above reasons, claims 1, 3-7, 9-12 and 16-19 are not anticipated by Chim, and Appellant respectfully requests that the rejection be reversed.

2. Are claims 13-15 unpatentable under 35 USC 103(a) over Chim, as applied to claim 7 above, and further in view of Steinberg?

Claims 13-15 are rejected under 35 USC 103(a) over Chim, as applied to claim 7 above, and further in view of Steinberg.

Although Chim does not disclose outputting the criteria for camera movement through a serial connection, a parallel connection or a network, Steinberg is cited to show such a teaching.

While not conceding the patentability per se of claims 13-15, it is urged that these claims are patentable by virtue of their dependency on claim 7.

Accordingly, the rejection of claims 13-15 under 35 USC 103(a) is in error and Appellant respectfully requests that the rejection should be reversed.

CONCLUSION

The rejection of claims 1, 3-7, 9-12 and 16-19 under 35 USC 102(e) as being anticipated by Chim (U.S. patent 6,275,258) and the rejection of claims 13-15 as being unpatentable under 35 USC 103(a) over Chim in view of Steinberg et al. (U.S. patent 6,750,902) are both in error for the reasons advanced above. Accordingly, Appellant respectfully requests that the

Board of Patent Appeals and Interferences reverse the rejections.

Respectfully submitted,

John C Fox

John C. Fox, Reg. 24,975
Consulting Patent Attorney
203-329-6584

APPENDIX

CLAIMS ON APPEAL

1. A method for automatically controlling the movements of at least one camera or camera lens to change the prospective of a scene viewed by said at least one camera or camera lens, said method comprising the steps of:

selecting at least one sequence of camera parametrics from a plurality of sequences of camera parametrics, wherein said at least one sequence of camera parametrics is selected from the group of camera movements including scanning, zooming, tilting, orientating, panning, fading, zoom-and-pull-back, fade-in, fade-out, and wherein said parametrics provide instruction to control movement of said at least one camera or camera lens;

determining criteria for executing said selected sequence of camera parametrics, wherein said criteria are responsive to at least one high level parameter of at least one object contained in said scene; and

adjusting movement of said at least one camera or camera lens in response to said determined criteria.

3. The method as recited in claim 1 wherein said at least one high level parameter includes the number of objects within said scene.

4. The method as recited in claim 1 wherein said at least one high level parameter includes the position of at least one object within said scene.

5. The method as recited in claim 1 wherein said at least one high level parameter includes speech recognition of at least one object within said scene.

6. The method as recited in claim 1 wherein said at least one high level parameter includes an audio input of at least one object within said scene.

7. An apparatus for automatically controlling the movements of at least one camera or camera lens to change the prospective of a scene viewed by said at least one camera or camera lens, said apparatus comprising:

a processor operative to:

receive a first input for selecting at least one sequence of camera parametrics from a plurality of sequences of camera parametrics, wherein said at least one sequence of camera parametrics is selected from the group of camera movements including scanning, zooming, tilting, orientating, panning, fading, zoom-and-pull-back, fade-in, fade-out, and wherein said parametrics provide instruction to control movement of said at least one camera or camera lens;

receive a second input comprising at least one high level parameter of at least one object contained in said scene;

determine criteria for executing said selected sequence of camera parametrics, wherein said criteria are responsive to said at least one high level parameter; and

means for adjusting movement of said at least one camera or camera lens in response to said determined criteria.

9. The apparatus as recited in claim 7 wherein said at least one high level parameter includes the number of objects within

said scene.

10. The apparatus as recited in claim 7 wherein said at least one high level parameter includes the position of at least one object within said scene.

11. The apparatus as recited in claim 7 wherein said at least one high level parameter includes speech recognition of at least one object within said scene.

12. The apparatus as recited in claim 7 wherein said at least one high level parameter includes an audio input of at least one object within said scene.

13. The apparatus as recited in claim 7 wherein said means for adjusting said camera movement effects outputting of said criteria over a serial connection.

14. The apparatus as recited in claim 7 wherein said means for adjusting said camera movement effects outputting of said criteria over a parallel connection.

15. The apparatus as recited in claim 7 wherein said means for adjusting said camera movement effects outputting of said criteria over a network.

16. The apparatus as recited in claim 7 wherein said camera movement is accomplished electronically.

17. The apparatus as recited in claim 7 wherein said camera movement is accomplished mechanically.

18. A method as in claim 1 including:

- locating the at least one object in an image of the scene;
- determining the object closest to a predetermined location in the image;
- adjusting the movement of the at least one camera or camera lens in response to said determination.

19. A method as in claim 1 including:

- locating the at least one object in an image of the scene;
- determining the object closest to the center of the image;
- determining the percentage of the scene around said closest object;
- adjusting the zoom level of the at least one camera or camera lens in response to said percentage determination.

EVIDENCE APPENDIX

(none)

RELATED PROCEEDINGS APPENDIX

(none)